Theoretical Computer Science 52 (1987) 59–75
North-Holland

ON REGULAR TRACE LANGUAGES *

Jacques SAKAROVITCH
Laboratoire d'Informatique Théorique et Programmation C.N.R.S. Paris,
Université de Paris VI, 75251 Paris, Cedex 05, France

Communicated by M. Nivat
Received November 1986

Abstract. We characterize here the free partially commutative monoids the regular sets of which form a Boolean algebra or are all unambiguous: these are, in both cases, the free products of free commutative monoids. This result has been established independently by other authors but the method used here is original. It is based on the properties of generalized automata on free products of monoids.

Résumé. Nous caractérisons ici les monoïdes de commutation dont les parties rationnelles forment une algèbre de Boole et ceux dont les parties rationnelles sont toutes non ambiguës: ce sont, dans les deux cas, les produits libres de monoïdes commutatifs libres. Ce résultat a été établi, indépendamment, par d'autres auteurs mais la méthode employée ici, qui utilise les propriétés des automates généralisés sur des produits libres de monoïdes, est originale.

1. Introduction

Since a long time already, the free partially commutative monoids have been considered in connection with combinatorial problems (cf. [7, 16]). More recently, words over a partially commutative alphabet became of interest to computer scientists for they allow to model problems of concurrency control. In this framework, the alphabet consists in functions and the commutation between these functions corresponds to the commutation of mappings under composition (cf. [23], for instance). Sets of words over such partially commutative alphabets were introduced by Mazurkiewicz under the name of *trace languages*, as a tool for describing the behaviour of concurrent program schemes in the same sense as usual formal languages describe the behaviour of sequential program schemes. Three recent surveys [2, 17, 20] give a rather complete description of the subject.

A paper by Flé and Roucairol [13] initiated a series of studies of trace languages that are recognizable [9, 10, 18], culminating in the characterization of recognizable trace languages by Ochmansky [19]. On the other hand, regular trace languages have become a natural object of study, as regular languages have within formal

* This work has been partially supported by the Programme de Recherche Coordonné 'Mathématiques et Informatique' of the Ministère de la Recherche et de la Technologie.

language theory. For instance, Bertoni, Mauri, and Sabadini considered in [5] the problems of membership and equivalence for regular trace languages. We address here, and we solve, the problem of characterizing the conditions on the commutation relations under which regular trace languages are closed under complementation and are unambiguously regular.

A few more formal definitions are convenient to present the content of this paper.

2. Results

Let A be a finite alphabet and let θ be a symmetrical relation on A which we shall call a commutation relation on A. We denote by $M(A, \theta)$ the quotient of the free monoid A^* by the congruence generated by the pairs of words (ab, ba) for all pairs (a, b) in θ. $M(A, \theta)$ is *the free partially commutative monoid* generated by A with respect to the relation θ; following Perrin [20], we call $M(A, \theta)$ a *commutation monoid*. If θ is empty, $M(A, \theta)$ is the free monoid A^* itself; if θ is the universal relation (i.e., $\theta = A \times A$), $M(A, \theta)$ is the free commutative monoid generated by A.

As in [1, 2, 5, 6, 17], subsets of $M(A, \theta)$ are called *trace languages*. Regular subsets of $M(A, \theta)$, that is, subsets of $M(A, \theta)$ defined by regular expressions, are called regular trace languages. Unambiguous regular expressions have been introduced in [12]; unambiguous regular subsets are those defined by unambiguous regular expressions. Their properties are studied in Eilenberg's treatise [11].

Strikingly enough, rational sets of commutation monoids share two remarkable properties in the two extremal situations: the free monoid and the free commutative monoid. From Kleene's theorem indeed, and from a part of its proof usually known as McNaughton and Yamada's algorithm, one deduces the following theorems.

Theorem A. *Rational sets of a finitely generated free monoid form a Boolean algebra (i.e. are closed under complementation).*

Theorem B. *Rational sets of a free monoid are unambiguously rational.*

On the other hand, the following results are now classical (as is shown in [12], the assumption of freeness in the two theorems is not necessary but this strong generalization plays no rôle here).

Theorem C (Ginsburg [15]). *Rational sets of a finitely generated free commutative monoid form a Boolean algebra.*

Theorem D (Eilenberg and Schützenberger [12]). *Rational sets of a free commutative monoid are unambiguously rational.*

We are here concerned with commutation monoids of a general type. An often-used example shows that the previous results cannot be extended without further hypothesis. The monoid $C=\{a, b\}^* \times \{c\}^*$ is a commutation monoid: $C=M(A, \theta)$ with $A=\{a, b, c\}$ and $\theta=\{(a, c), (c, a), (b, c), (c, b)\}$. Let P and Q be the subsets of C defined by

$$P=(a, c)^*(b, 1)^*, \qquad Q=(a, 1)^*(b, c)^*.$$

It is readily seen that

$$P \cap Q=\{(a^n b^n, c^n) \mid n \in \mathbb{N}\}$$

is not a regular subset of C and it belongs to folklore (cf. [11]) that $P \cup Q$ is an inherently ambiguous regular subset of C. We shall prove here the following theorem.

Theorem 2.1. *Let A be a finite alphabet and let $M=M(A, \theta)$ be a commutation monoid. The following three conditions are equivalent*:

(i) *the regular sets of M form a Boolean algebra*;

(ii) *the regular sets of M are unambiguously regular*;

(iii) *the relation θ is transitive.*

The above example gives the proof, ab absurdo, that conditions (i) or (ii) imply condition (iii). Indeed, if θ is not a transitive relation on A there exist three letters a, b, and c in A such that (a, c) and (c, b) are both in θ while (a, b) is *not* in θ. Thus, the monoid C is isomorphic to a submonoid of $M(A, \theta)$. Up to that isomorphism, the above sets P and Q are regular subsets of $M(A, \theta)$ and the conclusion follows.

After having written a first version of this paper we learned that Theorem 2.1 had been established independently by other authors. The implication (iii)$\Rightarrow$(i) is proved by Aalbersberg and Welzl in [1], and both (ii)$\Rightarrow$(i) and (iii)$\Rightarrow$(i) by Bertoni, Mauri, and Sabadini in [6]. The method used here however, and the results below from which we derive Theorem 2.1, appear to be new.

It is easily seen (cf. Section 3) that the relation θ on A is transitive if and only if $M(A, \theta)$ is isomorphic to a free product of free commutative monoids. Together with a proof of Theorem 2.1, the aim of this paper is to make it clear that this result is indeed the outcome of two independent types of results. Results of the first type are Theorems C and D stated above. In some sense, what commutativity brings to regular sets of a commutation monoid (as far as closure under complementation and ambiguity are concerned) is entirely taken into account by these two theorems. Results of the second type roughly say that the free product preserves closure under complementation and unambiguity of regular sets. More precisely, our main results are the two following theorems.

Theorem 2.2. *Let M and N be two monoids the regular subsets of which form a Boolean algebra. The regular subsets of the free product $M * N$ form a Boolean algebra.*

Theorem 2.3. *Let M and N be two monoids the regular subsets of which form a Boolean algebra and are unambiguously regular. The regular subsets of the free product $M * N$ are unambiguously regular.*

Clearly, Theorem C and Theorem 2.2 together give (iii)$\Rightarrow$(i) in Theorem 2.1; while Theorem C, Theorem D, and Theorem 2.3 together give (iii)$\Rightarrow$(ii). The paper is hence devoted to the proof of Theorems 2.2 and 2.3. Let us add a few remarks before getting to this core.

Remark 2.4. Recall first that two elements of a monoid M, different from the identity 1_M, are called *divisors of the identity* if their product is equal to 1_M. Recall also that a subset of a monoid M is said to be *recognizable* if it is the union of classes of a congruence of finite index of M (Kleene's theorem states that a subset of A^* is recognizable if and only if it is regular). We can now state a result of Reutenauer which bears a striking similarity with Theorems 2.2 and 2.3.

Theorem E (Reutenauer [21]). *Let M and N be two monoids without divisors of the identity and such that their regular subsets are recognizable. The regular subsets of the free product $M * N$ are recognizable.*

Indeed the method presented here allows to improve slightly that result as expressed in the following theorem.

Theorem 2.5. *Let M and N be two monoids the regular subsets of which are recognizable. The regular subsets of the free product $M * N$ are all recognizable subsets if and only if at least one of M or N has no divisors of the identity.*

It is true that commutation monoids have no divisors of the identity. But since not every regular subset of a free commutative monoid is recognizable, this last result has no connection with Theorem 2.1. Its proof is to be found elsewhere [22].

Remark 2.6. It is known that any finitely generated free group has the property that its regular subsets form a Boolean algebra and are unambiguously regular (cf. [3, 4, 14]). This may also be easily derived from Theorems 2.2 and 2.3: from the fact that $\mathbb{Z}$, the one-generator free group, is a subgroup of the free product $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$, one deduces that it has the required properties (cf. [22] for details); the conclusion follows since a finitely generated free group is the free product of a finite number of copies of $\mathbb{Z}$.

Remark 2.7. It is somewhat unsatisfactory that closure under complementation and unambiguity do not play independent and symmetrical rôles in Theorems 2.2 and 2.3 and that both properties are necessary in the hypothesis of Theorem 2.3. On the other hand, there is no example known of a monoid the regular subsets of which have one property and not the other. It is still an open problem to deduce one property from the other.

Remark 2.8. It can be noted that if not every regular subset of a commutation monoid $M(A, \theta)$ is unambiguous, $M(A, \theta)$ itself is always an unambiguous regular subset (of itself). This follows from the fact that for any commutation relation θ on an alphabet A it is possible to find a regular set of representatives of $M(A, \theta)$ in A^*, i.e., a regular cross-section of A^* for the congruence generated by θ (cf. [7] for the original result and [20] for further references).

A brief sketch of the proof of Theorems 2.2 and 2.3 will terminate this introductory section. We first give the definition of a *mechanism* on a monoid, a name which comes from Conway [8], but an idea that can be traced in the work of several others authors (cf. [11]). The basic result (which can be seen as Kleene's theorem for non-free monoids) is that mechanisms generate regular sets and that, conversely, any regular set can be generated by a mechanism (Theorem 4.2). Mechanisms are a handier tool to deal with regular sets than regular expressions. In a second step we consider *the alternating mechanisms* on a free product of two monoids and we shall see that any regular set of a free product can be generated by such an alternating mechanism (Proposition 5.1). Finally, the kernel of this work appears to be the proof that any regular set of a free product can be generated by a *proper alternating mechanism* (Theorem 5.2) and the study of the properties of such mechanisms (Propositions 6.1 and 6.2) from which the theorems we aim at are readily derived.

3. Free product

If M is a monoid, 1_M will denote its identity element.

The free product of two monoids M and N, denoted by $M * N$, can be identified with the monoid the elements of which are the finite sequences $(u_1, u_2, \ldots, u_n)$ of elements of $(M\backslash 1_M) \cup (N\backslash 1_N)$ alternatively taken in M and N, i.e.,

$$u_i \in M\backslash 1_M \Leftrightarrow u_{i+1} \in N\backslash 1_N,$$

where the sets $M\backslash 1_M$ and $N\backslash 1_N$ are supposed to be disjoint. The product of two such sequences

$$(u_1, u_2, \ldots, u_k)(v_1, v_2, \ldots, v_p)$$

is equal to

(i) $(u_1, u_2, \ldots, u_k, v_1, \ldots, v_p)$ if u_k and v_1 do not belong to the same monoid;

(ii) $(u_1, u_2, \ldots, u_{k-1}, u_k v_1, v_2, \ldots, v_p)$ if u_k and v_1 belong to the same monoid and $u_k v_1$ is different from the identity;

(iii) the product $(u_1, u_2, \ldots, u_{k-1})(v_2, v_3, \ldots, v_p)$ otherwise.

Both monoids M and N are submonoids of $M * N$; this allows us to simply write $u = u_1 u_2 \ldots u_k$ instead of $u = (u_1, u_2, \ldots, u_k)$. Such a factorization for an element u of $M * N$ is unique and is called its *canonical factorization*.

A *presentation* of a monoid M is a pair $\langle A;R\rangle$, where A is a set and R is a set of pairs of elements of A^* such that M is isomorphic to the quotient of A^* by the congruence generated by R. If $\langle A;R\rangle$ is a presentation of M and $\langle B;S\rangle$ is a presentation of N with A disjoint from B, then $\langle A\cup B;R\cup S\rangle$ is a presentation of the free product $M * N$.

A commutation relation θ on an alphabet A is transitive if and only if A can be partitioned into n disjoint subsets A_i: $A = A_1\cup A_2\cup\cdots\cup A_n$ such that θ is the union of the universal relation ω_i on each A_i. We have noted that the commutation monoid $M(A,\omega)$, where ω is the universal relation on A, is the free commutative monoid generated by A. If θ is transitive, $M(A,\theta)$ is hence a free product of free commutative monoids and, conversely, a free product of free commutative monoids is a commutation monoid $M(A,\theta)$ with a transitive commutation relation θ.

4. Mechanism on a monoid

Let us start from the classical definition of finite automata. A nondeterministic finite automaton is a quintuple $\mathfrak{A} = \langle Q, A, \delta, q, F\rangle$, where Q is the (finite) set of states, A the input alphabet, $\delta : A\times Q\to\mathcal{P}(Q)$ the transition function, q_0 the initial state, and F the set of final states. Mechanisms are a double generalization of finite automata. In the definition of $\delta : Q\times A\to\mathcal{P}(Q)$, let us interpret A as a set of elements of the free monoid A^*. Roughly speaking we have a mechanism in full generality if we replace, in the definition above, A^* by any monoid M and the set A by any set of subsets of M. We find it convenient to use mechanisms by means of their *matrix representation* for it allows shorter and easier proofs (even though they may then appear as less intuitive proofs); we shall thus give the formal definition of mechanisms under this representation.

Recall first that if M is a monoid, then the set of subsets of M, $\mathcal{P}(M)$, is canonically equipped with the multiplication:

$$\forall P, R\subset M \quad PR = \{pr \mid p\in P, r\in R\}.$$

Together with the union of sets, which will be denoted by "+" instead of "$\cup$", this multiplication gives $\mathcal{P}(M)$ the structure of a semiring. The subset $\{\emptyset, \{1_M\}\}$ of $\mathcal{P}(M)$ is a subsemiring isomorphic to the Boolean semiring $\mathbb{B} = \{0, 1\}$ and we shall freely use that identification. Since any sum, even infinite, is defined on $\mathcal{P}(M)$, the *star operator* is simply defined on $\mathcal{P}(M)$ by

$$\forall P\subset M \quad P^* = \bigcup_{n\in\mathbb{N}} P^n$$

with $P^0 = 1$ and $P^{n+1} = P^nP$ for every integer n.

If Q is any finite set, the $Q\times Q$-matrices (i.e., the matrices the rows and columns of which are indexed by Q) with entries in $\mathcal{P}(M)$ form a semiring denoted by $\mathcal{P}(M^{Q\times Q})$. We denote again by 1 the identity of $\mathcal{P}(M)^{Q\times Q}$, i.e., the identity

$Q \times Q$-matrix; the context should make clear to which identity 1 refers. The star operator is then defined on $\mathcal{P}(M)^{Q\times Q}$ as it was defined on $\mathcal{P}(M)$:

$$\forall X \in \mathcal{P}(M)^{Q\times Q} \quad X^* = \bigcup_{n\in\mathbb{N}} X^n$$

with $X^0 = 1$ and $X^{n+1} = X^n X$ for every integer n. We shall also use the notation $X^+ = \bigcup_{n\in\mathbb{N}_+} X^n$ and here we quote a few classical identities (cf. [8, 11]) which will be freely used in the sequel[1]

$$\forall Y, Z \in \mathcal{P}(M)^{Q\times Q} \quad \begin{aligned} &Y^* = 1 + Y^+,\\ &Y^+ = YY^* = Y^*Y,\\ &(Y+Z)^* = (Y^*Z)^*Y^* = Y^*(ZY^*)^*,\\ &(YZ)^+ = Y(ZY)^*Z,\\ &(Y^*)^* = Y^*. \end{aligned}$$

Before defining the mechanisms by their matrix representation, let us start again with the finite automata. Le $\mathfrak{A} = \langle Q, A, \delta, q_0, F\rangle$ be a finite automaton. Let X be the $Q \times Q$-matrix with entries in $\mathcal{P}(A)$ defined by

$$X_{p,q} = \{a \in A \mid q \in \delta(p, a)\}.$$

Let I be the Boolean row-vector of dimension Q defined by

$$I_q = \begin{cases} 1 & \text{if } q = q_0,\\ 0 & \text{otherwise;} \end{cases}$$

and let T be the Boolean column-vector of dimension Q defined by

$$T_p = \begin{cases} 1 & \text{if } p \in F,\\ 0 & \text{otherwise.} \end{cases}$$

The triple (I, X, T) is another description of the automaton $\mathfrak{A}$ and it is known that $L(\mathfrak{A})$, the language recognized by $\mathfrak{A}$, is equal to the set IX^*T (cf. [11], for instance).

Definition 4.1. Let M be a monoid and Q a finite set. *A mechanism* of dimension Q on M is a triple (I, X, T) where X is a $Q \times Q$-matrix the entries of which are subsets of M and where I is a Boolean row-vector and T a Boolean column-vector, both of dimension Q. The *resul* of a mechanism (I, X, T) is the subset IX^*T of M. Two mechanisms on M are *equivalent* if they have the same result. Finally, a mechanism (I, X, T) is said to be *proper* if the identity 1_M does not belong to any entry of X.

The matrix representation of finite automata goes back to the beginning of the theory of automata and may be considered as folklore. The name of 'mechanism' was introduced by Conway, who in [8], considered only mechanisms on free monoids

[1] All these identities but the last one are valid for Y and Z belonging to any semiring K where the star operation is defined; the last one holds because $Y + Y = Y$ for any Y in $\mathcal{P}(M)$.

and made a systematic use of the matrix representation. The mechanisms we have just defined are generalizations of the generalized M-automata in [11]. The following result is proved in [8] for M a free monoid; the same proof holds in the general case.

Theorem 4.2. *Let M be a monoid and $\mathscr{C}$ a family of subsets of M. A subset of M belongs to the regular closure of $\mathscr{C}$ if and only if it is the result of a mechanism the entries of which are in $\mathscr{C}$.*

The fact that finite automata with 'ε-moves' are equivalent to finite automata without ε-moves may be rephrased by the following lemma.

Lemma 4.3. *Any mechanism on M is equivalent to a proper mechanism on M.*

Proof. Let (I, X, T) be a mechanism on M. The matrix X may be written as $X = E + Y$ where E is a Boolean matrix and Y a matrix the entries of which do not contain 1_M. We then have

$$X^* = (E + Y)^* = (E^*Y)^*E^*.$$

The entries of E^*Y are a (finite) union of the entries of Y. The mechanism (I, E^*Y, E^*T) is thus proper and equivalent to (I, X, T). □

5. Alternating mechanism on a free product

Let M and N be two monoids. A mechanism (I, X, T) on the free product $M * N$ is called an *alternating mechanism* if there exists a block decomposition of X of the form

$$X = \begin{pmatrix} 0 & V \\ W & 0 \end{pmatrix}$$

such that the blocks in the diagonal are two null, square matrices and such that the entries of V belong to $\mathscr{P}(M)$ and the entries of W to $\mathscr{P}(N)$.

Proposition 5.1. *Any regular subset of $M * N$ is the result of an alternating mechanism the entries of which are regular subsets of M and N.*

Proof. From Theorem 4.2 it follows that any regular subset of $M * N$ is the result of a mechanism (I, X, T) the entries of which are finite subsets of $M + N$. One may thus write

$$X = Y + Z,$$

where Y is a mechanism on M and Z is a mechanism on N. The classical identities give

$$(Y+Z)^* = (Y^*Z)^*Y^* = (Z+Y^+Z)^*Y^* = Z^*(Y^+Z^+)^*Y^*$$
$$= (1+Z^+)(Y^+Z^+)^*(1+Y^+).$$

Developing and re-arranging this equality give

$$(Y+Z)^* = (Y^+Z^+)^* + Z^+(Y^+Z^+)^* + (Y^+Z^+)^*Y^+ + Z^+(Y^+Z^+)^*Y^*$$
$$= 1 + (Y^+Z^+)^+ + (Z^+Y^+)^*Z^+ + (Y^+Z^+)^*Y^+ + (Z^+Y^+)^+.$$

It is then not difficult to recognize that since $IT + IT = IT$, the mechanism (I, X, T) is equivalent to the alternating mechanism (I', X', T'), where

$$I' = (I \ I), \qquad X' = \begin{pmatrix} 0 & Y^+ \\ Z^+ & 0 \end{pmatrix}, \qquad T' = \begin{pmatrix} T \\ T \end{pmatrix}.$$

Since the entries of Y and Z are finite subsets of M and N, the entries of X' are regular subsets of M and N. □

In the sequel we shall say that an alternating mechanism on $M * N$ is *regular* if its entries are regular subsets of M and N. Lemma 4.3 implies that the mechanism (I, X, T) in the proof of Proposition 5.1 may always be assumed to be proper. If this holds and if neither M nor N have divisors of the identity (cf. Remark 2.4), then the regular alternating mechanism (I', X', T') *is also proper*. This is the case if M and N are free products of free commutative monoids. A reader who is only concerned with trace languages may thus jump to the next section where we shall derive Theorems 2.2 and 2.3 (and hence Theorem 2.1) from properties of proper alternating mechanisms.

If M and/or N have divisors of the identity, then Y^+ and/or Z^+ may well be nonproper even if Y and Z are proper. The key result of this section is that under rather a weak hypothesis every regular set of $M * N$ can nevertheless be generated by a proper alternating mechanism.

Theorem 5.2. *Let M be a monoid such that for every regular subset L of M, $L \setminus 1_M$ is also regular and let N be a monoid with the same property. Then any regular subset of $M * N$ is the result of a proper and regular alternating mechanism.*

Let us note that the hypothesis of Theorem 5.2 is obviously fulfilled when the regular subsets of M (respectively of N) form a Boolean algebra. (We already know that the conclusion of Theorem 5.2 holds if M and N have no divisors of the identity; it is an easy exercise to check if the hypothesis also holds in this case.) The proof of Theorem 5.2 relies on the following identities.

Lemma 5.3. *Let X be a matrix with a bloc decomposition*

$$X=\begin{pmatrix} 0 & V \\ E+Z & 0 \end{pmatrix}.$$

Then

(a) $$X^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 0 & (VE)^*V \\ Z & 0 \end{pmatrix}^*\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix};$$

(b) $$X^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 0 & (VE)^*V \\ E+Z & 0 \end{pmatrix}^*\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}.$$

Proof. Identity (a) is obtained by a sequence of computations on 2×2-matrices.

$$X^*=\begin{pmatrix} 0 & V \\ E+Z & 0 \end{pmatrix}^*=\left[\begin{pmatrix} 0 & 0 \\ E & 0 \end{pmatrix}+\begin{pmatrix} 0 & V \\ Z & 0 \end{pmatrix}\right]^*.$$

Since

$$\begin{pmatrix} 0 & 0 \\ E & 0 \end{pmatrix}^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix},$$

we have

$$X^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\left(\begin{pmatrix} 0 & V \\ Z & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\right)^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} VE & V \\ Z & 0 \end{pmatrix}^*$$

$$=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\left[\begin{pmatrix} VE & 0 \\ 0 & 0 \end{pmatrix}+\begin{pmatrix} 0 & V \\ Z & 0 \end{pmatrix}\right]^*.$$

Since

$$\begin{pmatrix} VE & 0 \\ 0 & 0 \end{pmatrix}^*=\begin{pmatrix} (VE)^* & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} (VE)^*VE & 0 \\ 0 & 0 \end{pmatrix},$$

we have

$$X^*=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\left(\begin{pmatrix} (VE)^* & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & V \\ Z & 0 \end{pmatrix}\right)^*\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} (VE)^*VE & 0 \\ 0 & 0 \end{pmatrix}\right)$$

$$=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 0 & (VE)^*V \\ Z & 0 \end{pmatrix}^*\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} 0 & (VE)^*V \\ Z & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ E & 0 \end{pmatrix}\right)$$

$$=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 0 & (VE)^*V \\ Z & 0 \end{pmatrix}^+\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}+\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}.$$

Since $E+E=E$, we now find

$$\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}$$

and (a) follows.

Let us develop the right-hand side of (b) using (a); we obtain

$$\begin{pmatrix} 0 & (VE)^*V \\ E+Z & 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}\begin{pmatrix} 0 & ((VE)^*VE)^*(VE)^*V \\ Z & 0 \end{pmatrix}^*\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}.$$

Since $((VE)^*VE)^*(VE)^* = (VE)^*$ and

$$\begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ E & 1 \end{pmatrix},$$

(b) follows, by using (a) the other way around. □

Exchanging the rows as well as the columns of X in Lemma 5.3 obviously gives (with the notation that will be used later):

$$\begin{pmatrix} 0 & F+Y \\ W & 0 \end{pmatrix}^* = \begin{pmatrix} 1 & F \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & Y \\ (WF)^*W & 0 \end{pmatrix}^* \begin{pmatrix} 1 & F \\ 0 & 1 \end{pmatrix},$$

an identity that will also be referred to as Lemma 5.3(a).

Proof of Theorem 5.2. (1) Let (I, X, T) be a regular alternating mechanism on $M * N$. Let

$$X = \begin{pmatrix} 0 & V_0 \\ W_0 & 0 \end{pmatrix}.$$

For every integer n, let V_n, W_n, E_n, Z_n, F_n, Y_n be inductively defined by the following formulae:

$$W_n = E_{n+1} + Z_{n+1},$$

where E_{n+1} is Boolean and Z_{n+1} proper;

$$V_{n+1} = (V_n E_{n+1})^* V_n = F_{n+1} + Y_{n+1},$$

where F_{n+1} is Boolean and Y_{n+1} proper;

$$W_{n+1} = (W_n F_{n+1})^* W_n.$$

The hypothesis on the regular sets of M and N implies that if X is regular, then so are Y_n and Z_n for every n. In order to simplify the notation, let us adopt the following conventions:

$g_0 = d_0 = 1$ identity matrix of the size of X;

$$e_n = \begin{pmatrix} 1 & 0 \\ E_n & 1 \end{pmatrix}, \qquad f_n = \begin{pmatrix} 1 & F_n \\ 0 & 1 \end{pmatrix};$$

$$g_{n+1} = g_n e_n f_n, \qquad d_{n+1} = f_n e_n d_n.$$

From Lemma 5.3 it follows that, for every n,

$$X^* = g_n \begin{pmatrix} 0 & V_n \\ W_n & 0 \end{pmatrix}^* d_n.$$

Let us remark that $E_{n+1} \geqslant E_n$ and $F_{n+1} \geqslant F_n$ for every n, from which follow $e_{n+1} \geqslant e_n$, $f_{n+1} \geqslant f_n$ as well as $g_{n+1} \geqslant g_n$ and $d_{n+1} \geqslant d_n$. These monotone sequences of Boolean matrices are constant from a certain rank k on. Let us note $E = E_k$, $F = F_k$, $e = e_k$, $f = f_k$, $g = g_k$, and $d = d_k$. One thus has

$$ge = g, \qquad gf = g, \qquad ed = d, \qquad fd = d.$$

From

$$V_k = (V_{k-1}E)^* V_{k-1} = F + Y_k,$$

$$V_{k+1} = (V_k E)^* V_k = ((V_{k-1}E)^* V_{k-1}E)^* V_{k-1} = V_k = F + Y_{k+1},$$

it follows that the sequence Y_n is constant from k onwards, and let thus $Y = Y_k$. The same is true for Z_n, so let $Z = Z_k$.

(2) Now, put

$$t = \begin{pmatrix} 0 & F \\ E & 0 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} 0 & Y \\ Z & 0 \end{pmatrix}.$$

One verifies that $t \leqslant ef$ and $t \leqslant fe$ from which $gt \leqslant g$ and $td \leqslant d$ follow. We now claim that

$$gx^*d = g(t+x)^*d \tag{1}$$

which we shall prove by showing by induction on n that

$$gx^{\leqslant n}d = g(t+x)^{\leqslant n}d$$

($x^{\leqslant n}$ is a shortening for $\sum_{i=0}^{i=n} x^i$). This equality clearly holds for $n = 0$ and $n = 1$.

We first note that

$$xtx = \begin{pmatrix} 0 & YEY \\ ZFZ & 0 \end{pmatrix}$$

and thus $xtx \leqslant (t+x)$. Now, $(t+x)^{n+1} = x^{n+1} + p(x, t)$, where p is a homogeneous polynomial of degree $n+1$ on the two noncommuting variables t and x. For every monomial $m(x, t)$ of p at least one of the following is true:

- $m(x, t) = tm'(x, t)$ where m' is of degree n
- $m(x, t) = m'(x, t)t$
- $m(x, t) = m'(x, t)xtxm'(s, t)$ where $m'm''$ is of degree $n-2$.

If we replace every factor x or t of m' and of m'' by $(t+x)$, we get

$$g(m(x, t))d \leqslant g(m'(x, t))d \leqslant g((t+x)^n)d,$$

$$g(m(x, t))d \leqslant g(m'(x, t)(t+x)m''(x, t))d \leqslant g((t+x)^{n-1})d.$$

Hence, $g(p(x, t))d \leqslant g((x+t)^{\leqslant n})d$, which proves (1) by induction hypothesis.

The mechanism (Ig, x, dT) is equivalent to (I, X, T), and proper. □

6. Properties of proper alternating mechanisms

The properties of proper alternating mechanisms are indeed deduced from the following observation: if P is a subset of M and if Q is subset of N, the products

PQ and QP are unambiguous; moreover, if both P and Q are proper (i.e., do not contain the identity), those subsets PQ and QP both generate free submonoids of $M * N$. In particular $M * N$ belongs to the unambiguous regular closure of $M^{\cdot} = M/1_M$ and $N^{\cdot} = N\backslash 1_N$:

$$M * N = (M^{\cdot}N^{\cdot})^* + (M^{\cdot}N^{\cdot})^* M^{\cdot} + N^{\cdot}(M^{\cdot}N^{\cdot})^* + N^{\cdot}(M^{\cdot}N^{\cdot})^* M^{\cdot}$$
$$= (1 + N^{\cdot})(M^{\cdot}N^{\cdot})^*(1 + M^{\cdot}). \tag{2}$$

Proposition 6.1. *Let L be the result of a proper alternating mechanism (E, X, T) on $M * N$. The complement of L in $M * N$ is the result of a (proper alternating) mechanism the entries of which belong to the Boolean algebra generated by the entries of X.*

Proof. (1) The first step is a generalization of the classical construction of a deterministic (finite) automaton from a nondeterministic one.

Let (E, V, T) be a proper mechanism on M of dimension Q: This mechanism defines a function $\delta : Q \times M \to \mathcal{P}(Q)$ by

$$\forall p \in Q\ \forall m \in M \quad \delta(p, m) = \{q \mid m \in V_{p,q}\}.$$

The function δ is additively extended to a function the domain of which is $\mathcal{P}(Q) \times M$ by

$$\forall I \in \mathcal{P}(Q) \quad \delta(I, m) = \bigcup_{p \in I} \delta(p, m).$$

In turn, this function δ allows to define a proper mechanism $\hat{V}$ of dimension $\mathcal{P}(Q)$:

$$\forall I, J \in \mathcal{P}(Q) \quad \hat{V}_{I,J} = \{m \in M^{\cdot} \mid \delta(I, m) = J\}.$$

The entries of $\hat{V}$ belong to the Boolean algebra generated by the entries of V. More precisely, the following claim holds.

Claim 1

$$\forall I, J \in \mathcal{P}(Q) \quad \hat{V}_{I,J} = \left[\bigcap_{q \in J}\left(\bigcup_{p \in I} V_{p,q}\right)\right] \Big\backslash \left[\bigcup_{q \notin J}\left(\bigcup_{p \in I} V_{p,q}\right)\right].$$

The proof of Claim 1 is given by the following sequence of equivalences:

$$\begin{aligned}
m \in \hat{V}_{I,J} &\Leftrightarrow \delta(I, m) = J \\
&\Leftrightarrow \{\forall q \in J, \exists p \in I, q \in \delta(p, m)\} \text{ and } \{\forall q \notin J, \forall p \in I, q \notin \delta(p, m)\} \\
&\Leftrightarrow \{\forall q \in J, \exists p \in I, q \in \delta(p, m)\} \text{ and } \neg\{\exists q \notin J, \exists p \in I, q \in \delta(p, m)\} \\
&\Leftrightarrow \{\forall q \in J, \exists p \in I, m \in V_{p,q}\} \text{ and } \neg\{\exists q \notin J, \exists p \in I, m \in V_{p,q}\} \\
&\Leftrightarrow \left\{m \in \bigcap_{q \in J}\left(\bigcup_{p \in I} V_{p,q}\right)\right\} \text{ and } \neg\left\{m \in \bigcup_{q \notin J}\left(\bigcup_{p \in I} V_{p,q}\right)\right\}.
\end{aligned}$$

The vectors $\hat{E}$ and $\hat{T}$ are defined in the same way, from E and T:

$$\hat{E}_J = \left[\bigcap_{q \in J} E_q\right] \Big\backslash \left[\bigcup_{q \notin J} E_q\right], \qquad \hat{T}_I = \bigcup_{p \in I} T_p.$$

Thus, since E and T are Boolean vectors,

$$\hat{E}_J = 1 \Leftrightarrow J = \{q \mid E_q = 1\}, \qquad \hat{T}_I = 1 \Leftrightarrow \exists p \in I\; T_p = 1.$$

At this point, the reader has to be aware that M plays the rôle of the alphabet A (and not of the monoid A^*) in the classical construction. For every I of $\mathcal{P}(Q)$ and for every m of M there exists one and only one J in $\mathcal{P}(Q)$ such that $m \in \hat{V}_{I,J}$. To that extent the mechanism $\hat{V}$ is deterministic and complete.

(2) Let now (E, X, T) be a proper alternating mechanism on $M * N$, and let

$$E = (K\ L), \qquad X = \begin{pmatrix} 0 & V \\ W & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} S \\ U \end{pmatrix}$$

be its alternating block decomposition. We suppose that both V and W are mechanisms of the same dimension Q. This causes no loss of generality for our purpose since we have in mind the construction of an alternating mechanism by means of Proposition 5.1. Let (E', X', T') be the mechanism obtained from (E, X, T) by applying the above construction to each of its components:

$$E' = (\hat{K}\ \hat{L}), \qquad X' = \begin{pmatrix} 0 & \hat{V} \\ \hat{W} & 0 \end{pmatrix}, \qquad T' = \begin{pmatrix} \hat{S} \\ \hat{U} \end{pmatrix}.$$

Claim 2. *(E', X', T') is equivalent to (E, X, T).*

Claim 3. *The result of $(E', X', \bar{T}')$ is the complement of the result of (E', X', T'), where $\bar{T}'$ is the complement of the Boolean vector T'.*

The proof of both claims exactly follows the proof of corresponding results on classical (finite) automata the input of which are words of a free monoid. The notation is slightly complicated by the fact that there are two functions δ and η (which correspond respectively to $\hat{V}$ and $\hat{W}$) that are applied alternatively instead of one.

Let $x = m_1 n_1 m_2 n_2 \ldots m_k n_k$ be an element of $M * N$ with its unique canonical factorization (there are in fact four cases to be considered, according to the decomposition of $M * N$ into four disjoint sets as in Eq. (2). The function $\delta * \eta : \mathcal{P}(Q) \times M * N \to \mathcal{P}(Q)$ is inductively defined on k by

$$\forall I \in \mathcal{P}(Q) \quad \delta * \eta(I, x) = \eta(\delta(\delta * \eta(I, m_1 n_1 m_2 n_2 \ldots m_{k-1} n_{k-1}), m_k), n_k).$$

One verifies, again by induction on k and as in the classical case, that

$$\forall I \in \mathcal{P}(Q) \quad \delta * \eta(I, x) = \{q \in Q \mid \exists p = p_1, q_1, p_2, q_2, \ldots, p_k, p_{k+1} = q,$$
$$p \in I\ \forall i, 1 \leq i \leq k, m_i \in V_{p_i, q_i}, n_i \in W_{q_i, p_{i+1}}\}$$

and the conclusion of both claims follows. □

Proof of Theorem 2.2. This proof is now straightforward: Let M and N be monoids the regular subsets of which form a Boolean algebra. In particular, they fulfill the hypothesis of Theorem 5.2. Let R be a regular set of $M * N$. By Theorem 5.2, R is the result of a proper and regular alternating mechanism (E, X, T). By Proposition 6.1, $\bar{R}$, the complement of R in $M * N$, is the result of a mechanism (E', X', T') the entries of which belong to the Boolean closure of the entries of X. The hypothesis on M and N imply that the entries of X' are again regular. The set $\bar{R}$ is then (by Theorem 4.2) regular. □

Proposition 6.2. *The result of a proper alternating mechanism (E, X, T) on $M * N$ belongs to the unambiguous regular closure of the Boolean closure of the entries of X.*

Proof. (1) Let (E, X, T) be a proper alternating mechanism on $M * N$:

$$E = (K \ L), \qquad X = \begin{pmatrix} 0 & V \\ W & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} S \\ U \end{pmatrix}$$

such that, as above, V and W have the same dimension Q. And, as above, let (E', X', T') be the mechanism obtained from (E, X, T) by 'determinization' of its components:

$$E' = (\hat{K} \ \hat{L}), \qquad X' = \begin{pmatrix} 0 & \hat{V} \\ \hat{W} & 0 \end{pmatrix}, \qquad T' = \begin{pmatrix} \hat{S} \\ \hat{U} \end{pmatrix}.$$

The next claim, another property of such determinized alternating mechanism, is again a generalization of a property of classical deterministic finite automata.

Claim 4. *The result of (E', X', T') is in the unambiguous regular closure of the entries of $\hat{V}$ and $\hat{W}$.*

Recall first that the result of a mechanism is in the unambiguous regular closure of its entries if every element of the result is obtained with a multiplicity equal to one (cf. [11] for details). Consider now the result of (E', X', T'). We have

$$X'^* = \begin{pmatrix} (\hat{V}\hat{W})^* & (\hat{V}\hat{W})^*\hat{V} \\ (\hat{W}\hat{V})^*\hat{W} & (\hat{W}\hat{V})^* \end{pmatrix}$$

and

$$E'X'^*T' = \hat{K}(\hat{V}\hat{W})^*\hat{S} + \hat{K}(\hat{V}\hat{W})^*\hat{V}\hat{U} + \hat{L}(\hat{W}\hat{V})^*\hat{W}\hat{S} + \hat{L}(\hat{W}\hat{V})^*\hat{U}. \tag{3}$$

Since the entries of $\hat{V}$ and of $\hat{W}$ are all proper, the sum in (3) is unambiguous by the remark we made at the beginning of this section.

Let $x = m_1 n_1 m_2 n_2 \ldots m_k n_k$ be an element of the result of (E', X', T'). Then, necessarily, x belongs to $\hat{K}(\hat{V}\hat{W})^*\hat{S}$ and, even more, x necessarily belongs to $\hat{K}(\hat{V}\hat{W})^k\hat{S}$. By the definition of $\hat{K}$, there is a unique subset I of $\mathcal{P}(Q)$ such that $\hat{K}_I \neq 0$; by the definition of $\hat{V}$, there is a unique J such that m_1 belongs to $\hat{V}_{I,J}$; and, by the definition of $\hat{W}$, there is a unique F such that n_1 belongs to $\hat{W}_{J,F}$. Thus, and by induction on k, the element x is obtained with multiplicity one in the result of (E', X', T'). The three other cases are quite the same, and Claim 4 is therefore proved.

Since the entries of $\hat{V}$ and $\hat{W}$ belong to the Boolean closure of the entries of V and W, the proposition is established. □

Proof of Theorem 2.3. This now goes exactly as the one of Theorem 2.2: Let M and N be two monoids the regular subsets of which form a Boolean algebra and are unambiguously regular; in particular, they fulfill the hypothesis of Theorem 5.2. Let R be a regular set of $M * N$. By Theorem 5.2, R is the result of a proper and regular alternating mechanism (E, X, T). By Proposition 6.2, R belongs to the unambiguous regular closure of the Boolean algebra generated by the entries of X. The hypothesis on M and N imply that every element of that Boolean algebra is an unambiguous regular set of $M * N$: R is thus an unambiguous regular set of $M * N$. □

References

[1] IJ.J. Aalbersberg and E. Welzl, Trace languages defined by regular string languages, *RAIRO Inform. Théor.*, to appear.

[2] IJ.J. Aalbersberg and G. Rozenberg, Theory of traces, Tech. Rept. 85-16, Dept of Computer Science, Univ. of Leiden, 1985.

[3] M. Benois, Parties rationnelles du groupe libre, *C.R. Acad. Sci. Paris Sér. A* **269** (1969) 1188-1190.

[4] J. Berstel, *Transductions and Context-free Languages* (Teubner, Stuttgart, 1979).

[5] A. Bertoni, G. Mauri and N. Sabadini, Equivalence and membership problems for regular trace languages, *Proc. 9th Internat. Conf. on Automata, Languages and Programming*, Lecture Notes in Computer Science **140** (Springer, Berlin, 1982) 61-71.

[6] A. Bertoni, G. Mauri and N. Sabadini, Unambiguous regular trace languages, in: J. Demetrovics, G. Katona and A. Salomaa, eds, *Algebra, Combinatorics, and Logic in Computer Science*, Colloquia Mathematica Societas Janos Bolyai **42** (North-Holland, Amsterdam, 1985).

[7] P. Cartier and D. Foata, *Problèmes Combinatoires de Commutations et Réarrangements*, Lecture Notes in Mathematics **85** (Springer, Berlin, 1969).

[8] J.H. Conway, *Regular Algebra and Finite Machines* (Chapman and Hall, London, 1971).

[9] R. Cori and D. Perrin, Automates et commutations partielles, *RAIRO Inform. Théor.* **19** (1985) 21-32.

[10] R. Cori and Y. Métivier, Recognizable subsets of some partially abelian monoids, *Theoret. Comput. Sci.* **35** (1985) 179-189.

[11] S. Eilenberg, *Automata, Languages and Machines, Vol. A* (Academic Press, New York/London, 1974).

[12] S. Eilenberg and M.P. Schützenberger, Rational sets in commutative monoids, *J. Algebra* **13** (1969) 344-353.

[13] M.P. Flé and G. Roucairol, Maximal serializability of iterated transactions, *Theoret. Comput. Sci.* **38** (1985) 1-16.

[14] M. Fliess, Deux applications de la représentation matricielle d'une série rationnelle non commutative, *J. Algebra* **19** (1971) 344-353.

[15] S. Ginsburg and E.H. Spanier, Bounded ALGOL-like languages, *Trans. Amer. Math. Soc.* **113** (1964) 333-368.

[16] M. Lothaire, *Combinatorics on Words*, Encyclopedia of Mathematics **17** (Addison-Wesley, Reading, MA, 1983).

[17] A. Mazurkiewicz, Traces, histories, and graphs: instances of a process monoid, in: *Proc. of MFCS'84*, Lecture Notes in Computer Science **176** (Springer, Berlin, 1984) 115-133.

[18] Y. Metivier, Une condition suffisante de reconnaissabilité dans un monoïde partiellement commutatif, *RAIRO Inform. Théor.* **20** (1986) 121-127.

[19] E. Ochmanski, Regular behaviour of concurrent systems, *EATCS Bulletin 27*, October 1985.

[20] D. Perrin, Words over partially commutative alphabet, in: A Apostolico and Z. Galil, eds., *Combinatorial Algorithms on Words*, NATO-ASI Series (Springer, Berlin, 1985) 329-340.
[21] Ch. Reutenauer, Sur les semi-groupes vérifiant le théorème de Kleene, *RAIRO Inform. Théor.* **19** (1985) 281-291.
[22] J. Sakarovitch, Sur les parties rationnelles d'un produit libre, to appear.
[23] J. Ullman, *Principle of Database Systems* (Computer Science Press, Rockville, MD, 1980).